

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-160917

(43)Date of publication of application : 20.06.1997

(51)Int.Cl.

G06F 17/27
G06F 17/21

(21)Application number : 07-345264

(71)Applicant : DAINIPPON SCREEN MFG CO
LTD

(22)Date of filing : 06.12.1995

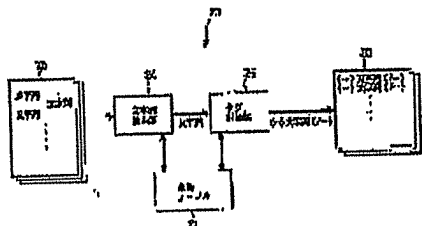
(72)Inventor : KADOMA HISAAKI

(54) PREPARING DEVICE FOR DOCUMENT WITH TAG AND METHOD THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To easily prepare a document with tag even if a document preparing person does not understand the meaning of the tag, how to use the tag and the logical structure of a document.

SOLUTION: A conversion table 22 stores coordinates in a character input area in a routine format and a tag to be added to the character string existing at the coordinate location by making the coordinate and the tag correspond to each other. A character string extraction part 24 refers to the conversion table 22 and extracts the character string from the character input area of the routine document 28 prepared by using the routine format. A tag addition part 26 refers to the conversion table 26, adds the tag corresponding to the coordinates in the character input area where the character string is extracted before and after the extracted character string and outputs the obtained character string with tag as the part of a document 30 with tag.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-160917

(43) 公開日 平成9年(1997)6月20日

(51) Int.Cl. ⁹	識別記号	序内整理番号	F I	技術表示箇所
G 0 6 F 17/27			C 0 6 F 15/20	5 5 0 E
17/21				5 3 8 A

審査請求 未請求 請求項の数6 F D (全 9 頁)

(21) 出願番号 特願平7-345264

(22) 出願日 平成7年(1995)12月6日

(71) 出願人 00020/551

大日本スクリーン製造株式会社

京都府京都市上京区堀川通寺之内上る4丁目天神北町1番地の1

(72) 発明者 角間 央章

京都市上京区堀川通寺之内上る4丁目天神北町1番地の1 大日本スクリーン製造株式会社内

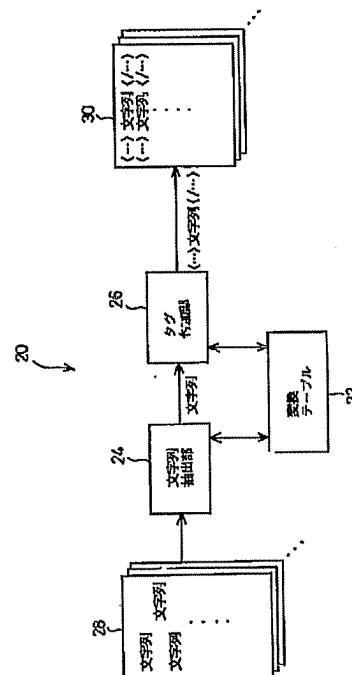
(74) 代理人 弁理士 五十嵐 孝雄 (外3名)

(54) 【発明の名称】 タグ付き文書作成装置及びその方法

(57) 【要約】

【課題】 文書作成者がタグの意味や使い方或いは文書の論理構造などを理解していなくても、容易にタグ付き文書を作成できるようにする。

【解決手段】 変換テーブル22は定型フォーマットにおける文字入力エリアの座標と、その座標位置にある文字列に付加すべきタグとを対応付けて格納している。文字列抽出部24は変換テーブル22を参照して、定型フォーマットを用いて作成された定型文書28の文字入力エリア内より、文字列を抽出する。タグ付加部26は変換テーブル22を参照して、文字列を抽出した文字入力エリアの座標に対応するタグを、抽出した文字列の前後に付加し、得られたタグ付きの文字列をタグ付き文書30の一部として出力する。



【特許請求の範囲】

【請求項1】 定型フォーマットの文字入力エリアに文字列を入力することにより作成された定型文書の中から、前記文字入力エリア内の文字列を抽出し、該文字列にタグを付加して、タグ付き文書を作成するタグ付き文書作成装置であって、前記文字入力エリアに関連する情報と、該文字入力エリア内より抽出した文字列に付加すべきタグとを対応付けて格納する変換テーブルと、前記文字入力エリア内より文字列を抽出する文字列抽出手段と、前記文字列を抽出した前記文字入力エリアの関連情報に対応するタグを、前記変換テーブルより読み出して、抽出した前記文字列に付加し、得られたタグ付きの文字列によって前記タグ付き文書を作成するタグ付加手段と、を備えるタグ付き文書作成装置。

【請求項2】 請求項1に記載のタグ付き文書作成装置において、前記文字列抽出手段は、前記変換テーブルより前記関連情報を読み出して、該関連情報に基づいて前記文字入力エリア内より文字列を抽出することを特徴とするタグ付き文書作成装置。

【請求項3】 請求項1または2に記載のタグ付き文書作成装置において、前記関連情報は、前記文字入力エリアの前記定型文書内での位置を表す位置情報を含むことを特徴とするタグ付き文書作成装置。

【請求項4】 請求項1または2に記載のタグ付き文書作成装置において、前記関連情報は、前記文字入力エリアの属性を表す属性情報を含むことを特徴とするタグ付き文書作成装置。

【請求項5】 タグ付き文書を作成するためのタグ付き文書作成方法であって、定型フォーマットの文字入力エリアに文字列を入力することにより作成された定型文書を用意する第1の工程と、前記文字入力エリアに関連する情報と、該文字入力エリア内より抽出する文字列に付加すべきタグとを対応付けて格納する変換テーブルを用意する第2の工程と、前記定型文書の中から、前記文字入力エリア内の文字列を抽出する第3の工程と、前記文字列を抽出した前記文字入力エリアの関連情報に対応するタグを、前記変換テーブルより読み出して、抽出した前記文字列に付加し、得られたタグ付きの文字列によって前記タグ付き文書を作成する第4の工程と、を備えるタグ付き文書作成方法。

【請求項6】 請求項5に記載のタグ付き文書作成方法において、前記第1の工程は、前記定型フォーマットを作成する工程を含むタグ付き文

書作成方法。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】本発明は、定型フォーマットを用いて作成された定型文書からタグ付き文書を作成するタグ付き文書作成装置及びその方法に関するものである。

【0002】

【従来の技術】タグ付き文書として代表的なものには、ISO（国際標準化機構）によるSGML（Standard Generalized Markup Language）規格がある。この規格によると、タグ付き文書とは図8に示すような文書を言う。即ち、タグ付き文書では、例えば<title>×××××××</title>のように、文字列（この場合、×××××××）の前後に、その文字列の内容、特徴、性質、性格などを表す情報（この場合、title）を含むタグが付加されている。

【0003】このようなタグ付き文書は、文書の蓄積や再利用、或いは文書の加工などを行なう場合にとても有効である。例えば、このようなタグ付き文書を多数蓄積した場合、タグを目印として、各文書の共通の情報（例えば、タイトル）を容易に検索抽出することができる。また、タグ付き文書を印刷する場合でも、特定のタグの付加された文字列のみを特殊な文字で印刷するなどの指定を行なうことができる。

【0004】

【発明が解決しようとする課題】しかし、このようなタグ付き文書を作成するためには、文書作成者が、タグの意味や使い方或いは文書の論理構造などを理解した上で、適切な位置にタグを挿入して文書を作成する必要があった。従って、一般の文書作成者がタグ付き文書を作成することは、大変手間のかかる作業となり、そのため、一般の文書の作成者によるタグ付け文書の作成は活発には行なわれていなかった。

【0005】従って、本発明の目的は、上記した従来技術の問題点を解決し、文書作成者がタグの意味や使い方或いは文書の論理構造などを理解していなくても、容易にタグ付き文書を作成することができるタグ付き文書作成装置を提供することにある。

【0006】

【課題を解決するための手段およびその作用・効果】上記した目的の少なくとも一部を達成するために、第1の発明は、定型フォーマットの文字入力エリアに文字列を入力することにより作成された定型文書の中から、前記文字入力エリア内の文字列を抽出し、該文字列にタグを付加して、タグ付き文書を作成するタグ付き文書作成装置であって、前記文字入力エリアに関連する情報と、該文字入力エリア内より抽出した文字列に付加すべきタグとを対応付けて格納する変換テーブルと、前記文字入力エリア内より文字列を抽出する文字列抽出手段と、前記

文字列を抽出した前記文字入力エリアの関連情報に対応するタグを、前記変換テーブルより読み出して、抽出した前記文字列に付加し、得られたタグ付きの文字列によって前記タグ付き文書を作成するタグ付加手段と、を備えることを要旨とする。

【0007】このように、第1の発明では、定型フォーマットの文字入力エリアに文字列を入力することによって作成された定型文書から、タグ付き文書を作成している。即ち、文字列抽出手段が、作成された定型文書の文字入力エリア内より文字列を抽出し、タグ付加手段が、文字列を抽出した文字入力エリアの関連情報に対応するタグを、変換テーブルより読み出して、抽出した文字列に付加し、得られたタグ付きの文字列によってタグ付き文書を作成する。なお、変換テーブルは、定型フォーマットにおける各文字入力エリアの関連情報と、その文字入力エリア内より抽出した文字列に付加すべきタグとを対応付けて格納している。

【0008】従って、第1の発明によれば、文書作成者がタグの意味や使い方或いは文書の論理構造などを理解していなくても、定型フォーマットを用いて定型文書を作成しさえすれば、その定型文書から容易にタグ付き文書を作成することができる。

【0009】第1の発明のタグ付き文書作成装置において、前記文字列抽出手段は、前記変換テーブルより前記関連情報を読み出して、該関連情報に基づいて前記文字入力エリア内より文字列を抽出するようにしても良い。

【0010】変換テーブルには文字入力エリアの関連情報が格納されているので、その関連情報を用いて文字入力エリアを特定し、その入力エリア内より文字列を抽出することができる。そして、その抽出した文字列に読み出した上記関連情報に対応するタグを付加すれば良い。

【0011】第1の発明のタグ付き文書作成装置において、前記関連情報は、前記文字入力エリアの前記定型文書内での位置を表す位置情報を含むことが好ましい。また、前記文字入力エリアの属性を表す属性情報を含むことが好ましい。これらの情報は、いずれも、これらの情報がわかれば、関連する文字入力エリアを特定することができる情報だからである。

【0012】第2の発明は、タグ付き文書を作成するためのタグ付き文書作成方法であって、定型フォーマットの文字入力エリアに文字列を入力することにより作成された定型文書を用意する第1の工程と、前記文字入力エリアに関連する情報と、該文字入力エリア内より抽出する文字列に付加すべきタグとを対応付けて格納する変換テーブルを用意する第2の工程と、前記定型文書の中から、前記文字入力エリア内の文字列を抽出する第3の工程と、前記文字列を抽出した前記文字入力エリアの関連情報に対応するタグを、前記変換テーブルより読み出して、抽出した前記文字列に付加し、得られたタグ付きの文字列によって前記タグ付き文書を作成する第4の工程

と、を備えることを要旨とする。

【0013】このように、第2の発明では、第1の工程において、定型フォーマットを用いて作成された定型文書を用意し、第2の工程において、文字入力エリアの関連情報とタグとを対応付ける変換テーブルを用意する。そして、第3の工程では、定型文書の文字入力エリア内より文字列を抽出し、第4の工程では、文字列を抽出した文字入力エリアの関連情報に対応するタグを変換テーブルより読み出して、文字列に付加し、タグ付き文書を作成する。

【0014】従って、第2の発明においても、第1の発明と同様に、文書作成者がタグの意味や使い方或いは文書の論理構造などを理解していなくても、容易にタグ付き文書を作成することができる。

【0015】第2の発明のタグ付き文書作成方法において、前記第1の工程は、前記定型フォーマットを作成する工程を含んでも良い。

【0016】

【発明の実施の形態】以下、本発明の実施の形態を実施例に基づいて説明する。図1は本発明の第1の実施例としてのタグ付き文書作成装置を示すブロック図である。図1に示すように、本実施例のタグ付き文書作成装置20は、定型フォーマット内の座標とその座標位置にある文字列に付加すべきタグとを対応付けて格納する変換テーブル22と、変換テーブル22を参照して、定型フォーマットを用いて作成された定型文書より文字列を抽出する文字列抽出部24と、変換テーブル22を参照して、抽出された文字列の前後にタグを付加するタグ付加部26と、を備えている。

【0017】なお、図1に示す装置によってタグ付き文書を作成するには、図2に示すように、前もって、定型フォーマットの作成（ステップS10）や、変換テーブルの作成（ステップS12）や、定型フォーマットを用いた文書入力（ステップS14）などの各処理を行なう必要がある。以下、図2に従って、順に説明する。

【0018】まず、文書を入力する際に用いる定型フォーマットを作成する処理を行なう（ステップS10）。図3は作成された定型フォーマットの一例を示す説明図である。この例は「技術報告書」の定型フォーマットの例である。図3に示すように、各記入項目にはそれぞれ文字入力エリアがそれぞれ設定されている。例えば、エリアAは発行日を入力するためのエリアであり、エリアBは件名を、エリアCは期間を、エリアDは技術分類を、エリアEはキーワードを、それぞれ入力するためのエリアである。作成された定型フォーマットはデータとしてハードディスクなどの記憶手段（図示せず）に蓄積される。なお、定型フォーマットについては、自ら作成しても良いし、また、市販されている汎用の定型フォーマットを利用するようにしても良い。

【0019】次に、図1の変換テーブル22の内容を作

成する処理を行なう(ステップS12)。即ち、予め、定型フォーマットを作成する際に各文字入力エリアA～Eの座標を調べておき、各座標に対して、各文字入力エリアA～Eに入力される文字列に関連するタグをそれぞれ設定して、変換テーブル22の内容を作成する。

【0020】図4は図3の定型フォーマットにおける各文字入力エリアA～Eの座標を示す説明図であり、また、図5は作成された変換テーブルの内容の一例を示す説明図である。図4に示すように、各文字入力エリアA～Eの座標としては、例えば、各エリアの先頭の左下点の座標を用いる。従って、文字入力エリアAの座標は例えば(146, 25)となり、文字入力エリアBの座標は(16, 41)となる。一方、文字入力エリアAには文字列として発行日に関する文字列の入力が予定されているので、この文字列に付加すべきタグとして<date></date>を設定する。また、文字入力エリアBには件名に関する文字列の入力が予定されているので、この文字列に付加すべきタグとしては<title></title>を設定する。従って、変換テーブル22の内容として、図5に示すように、座標(146, 25)にはタグとして<date></date>が、座標(16, 41)にはタグとして<title></title>が、それぞれ対応付けられる。こうして作成された変換テーブル22の内容は、変換テーブル22内に格納される。

【0021】なお、このような定型フォーマットと変換テーブルの作成は、将来における文書の蓄積や再利用、或いは文書の加工などが行い易くなるように、文書の論理構造を良く理解し、タグの意味や使い方を良く知っている者によって、それぞれ行なわれる。

【0022】さて、こうして定型フォーマットと変換テーブルが作成されたら、次に、定型フォーマットを用いて文書を入力する処理を行なう(ステップS14)。文書の入力、各文字入力エリアA～Eにそれぞれ所望の文字列を入力することによって行なわれる。図6は図3の定型フォーマットを用いて文書を入力した状態を示す説明図である。図6に示すように、各文字入力エリアA～Eに文字列が入力されると、それら文字入力エリアA～Eは、入力される文字列の長さに応じて、自動的にエリアの長さが拡大される。

【0023】こうした文書の入力、文書の論理構造の知識がない一般の文書作成者によって行なわれることを想定している。即ち、文書作成者は定型フォーマットにおいて準備されている各記入項目へ所望の内容を入力するだけで良いため、一般の文書作成者でも容易に文書入力を行なうことができる。なお、こうして定型フォーマットを用いて作成された定型文書(即ち、入力された文書)は、ハードディスクなどの記憶手段(図示せず)に記憶される。

【0024】以上のようにして、定型フォーマットを用いて作成された定型文書が多数蓄積されたら、次に、図

1に示したタグ付き文書作成装置によって、タグ付き文書を作成する処理を行なう(ステップS16)。図1に示すように、タグ付き文書30は、定型フォーマットを用いて作成された定型文書28の一つ一つに対して、一つずつ作成される。

【0025】図7は図2のタグ付き文書作成ステップS16の詳細処理ルーチンを示すフローチャートである。なお、図7に示す処理ルーチンは、定型フォーマットを用いて作成された一つの定型文書28について、一つのタグ付き文書30を作成しようとする度に起動される。

【0026】図7に示すように、まず、この処理ルーチンが起動されると、変換テーブル22より、格納された座標を読み出す処理を行なう(ステップS20)。即ち、図1の文字列抽出部24が、変換テーブル22に格納された複数の座標のうち、先頭に格納されている座標を一つ読み出す。従って、変換テーブル22の内容が図5に示した如くであるとする、座標として(146, 25)がまず読み出されることになる。

【0027】次に、定型フォーマットを用いて作成された定型文書28から、読み出した座標位置にある文字列を抽出する処理を行なう(ステップS22)。即ち、文字列抽出部24が、読み出した座標位置の文字入力エリア内に入力されている文字列を抽出する。従って、作成された定型文書28が図6に示した如くであり、読み出した座標が(146, 25)であるとする、座標(146, 25)に対応する文字入力エリアはAであるので、そのエリアAに入力されている文字列「〇年〇月〇日」が抽出される。なお、エリアA内には「発行日:」という文字列も存在するが、この文字列はタグ付き文書を構成する情報として特に必要はないので、抽出の対象から除外される。但し、「発行日:」という文字列も含めて抽出するようにしても、何ら支障はない。

【0028】一般に、市販されているDTP(ディスク・トップ・パブリッシング)ソフトなどでは、文書内での文字列や画像の入力位置をその座標で記憶するようにしており、さらに、座標を指定すると、その座標位置にある文字列や画像を読み出すことができるようになっていものが多い。従って、文字列抽出部24は、そのようなソフトを利用することにより、ステップS22の処理を実現することが可能である。なお、自らDTPソフトを作成して利用する場合でも、上記したような市販のDTPソフトと同等の機能があれば十分である。

【0029】次に、変換テーブル22を参照して、抽出した文字列の前後に、読み出した座標に対応するタグを付加する処理を行なう(ステップS24)。即ち、タグ付加部26が、ステップS20で読み出した座標に対応するタグを変換テーブル22より読み出して、抽出した文字列の前後にそのタグを付加する。こうして得られたタグ付きの文字列はタグ付き文書30の一部として出力される。

【0030】従って、変換テーブル22の内容が図5に示した如くであり、座標が(146, 25)である場合、タグとしては「<date></date>」が読み出され、ステップ22で抽出された文字列「〇年〇月〇日」の前後に付加されて、タグ付きの文字列「<date>〇年〇月〇日</date>」が得られ、これがタグ付き文書の一部となる。

【0031】次に、変換テーブル22に読み出すべき座標が残っているか否かを判定する処理を行なう(ステップS26)。即ち、文字列抽出部24が、変換テーブル22を再び参照して、変換テーブル22に格納されている座標のうち、未だ読み出されていない座標が存在するかどうかを判定する。そして、存在する場合にはステップ20に戻って次の座標を読み出す処理を行ない、以下、前述と同様の処理を繰り返す。存在しない場合には、図7に示す処理ルーチンを終了する。処理ルーチンを終了すると、定型フォーマットを用いて作成された一つの定型文書28について、一つのタグ付き文書30が作成されたことになる。

【0032】図8は図6の定型文書について作成されたタグ付き文書の一例を示す説明図である。図5に示す変換テーブル22の内容に従って、図6の定型文書内の所定の文字列にタグが付加されることによって、図8に示すようなタグ付き文書が作成される。

【0033】このようにして、定型フォーマットを用いて作成された一つの定型文書28について、一つのタグ付き文書30が作成されたら、蓄積された次の定型文書についてタグ付き文書を作成するために、再び、図7の処理ルーチンを起動する。

【0034】以上述べたように、本実施例によれば、文書作成者はタグの意味や使い方や文書の論理構造などを理解していなくても、単に定型フォーマットを用いて文書を入力するだけで良く、後は、図1に示したタグ付き文書作成装置によって自動的にタグ付き文書を作成することができるので、タグ付き文書の作成に手間がかからない。

【0035】次に、本発明の第2の実施例について説明する。本実施例の構成は、図1に示した第1の実施例の構成と同じであり、本実施例における処理手順も図2に示した第1の実施例の処理手順と同じである。但し、変換テーブル22の内容が第1の実施例の場合と異なる。本実施例では、変換テーブル22の内容として文字入力エリアの座標を用いるのではなく、文字入力エリアの属性を用いるようにする。

【0036】即ち、本実施例では、予め、定型フォーマットを作成する際に、各文字入力エリアA～Eに対してそれぞれ属性を設定しておき、各属性に対して、各文字入力エリアA～Eに入力される文字列に関連するタグをそれぞれ設定して、変換テーブル22の内容を作成する。一般に、市販されているDTPソフトなどでは、文

書内での文字入力エリアに対して文字フォントや文字サイズの他、属性を設定できる機能を有するものがあり、ここでは、そのような機能を有するソフトを用いることによって上記処理を実現することができる。

【0037】図9は図3の定型フォーマットにおける各文字入力エリアA～Eに対して設定された属性の一例を示す説明図であり、また、図10は作成された変換テーブルの内容の一例を示す説明図である。図9に示すように、各文字入力エリアA～Eに対する属性としては、各記入項目のタイトルをそのまま用いる。従って、文字入力エリアAに対する属性は「発行日」となり、文字入力エリアBに対する属性は「件名」となる。一方、文字入力エリアAの文字列に付加すべきタグとしては、第1の実施例と同様に、<date></date>を設定し、また、文字入力エリアBの文字列に付加すべきタグとしては、<title></title>を設定する。従って、変換テーブル22の内容としては、図10に示すように、属性「発行日」にはタグとして<date></date>が、属性「件名」にはタグとして<title></title>が、それぞれ対応付けられる。こうして作成された変換テーブル22の内容は、変換テーブル22内に格納される。

【0038】なお、ここで、属性とは文字入力エリアを識別するための指標となるものをいい、従って、文字入力エリアを識別することができるものであれば、単なる番号であっても、記号であっても良く、定型フォーマットを作成する者が自由に設定することができる。また、異なる文字入力エリアにそれぞれ入力される文字列に対し、同一のタグを付加したいのであれば、それら文字入力エリアに対して属性を設定する際に、それぞれ同一の属性を設定するようにすれば良い。例えば、図9の文字入力エリアAとCに入力される文字列に、それぞれ、<date></date>というタグを付加したいのであれば、それぞれのエリアA、Cに「発行日」という属性を設定すれば良い。

【0039】さて、このようにして内容の作成された変換テーブル22を用いて、タグ付き文書の作成を行なう場合の処理について、次に説明する。図11は本発明の第2の実施例におけるタグ付き文書作成ステップS16の詳細処理ルーチンを示すフローチャートである。なお、図11に示す処理ルーチンについても、図7の処理ルーチンと同様に、定型フォーマットを用いて作成された一つの定型文書28について、一つのタグ付き文書30を作成しようする度に起動される。

【0040】図11に示すように、まず、この処理ルーチンが起動されると、作成された定型文書28の文字入力エリアより、文字列を抽出する処理を行なう(ステップS30)。即ち、文字列抽出部24が、作成された定型文書28における複数の文字入力エリアのうち、最初の文字入力エリア内に入力されている文字列を抽出する。従って、作成された定型文書28が図6に示した如

くである場合、最初の文字入力エリアはAであるので、そのエリアAに入力されている文字列「〇年〇月〇日」が抽出される。なお、抽出された文字列から「発行日：」という文字列が除外されているのは、第1の実施例と同様の理由による。

【0041】次に、文字列を抽出した文字入力エリアの属性を調べる処理を行なう（ステップS32）。即ち、文字列抽出部24が、文字列を抽出した文字入力エリアに対して設定されている属性が何であるかを調べる。従って、例えば、各文字入力エリアA～Eに対して属性が図9に示す如く設定されている場合、文字列を抽出した文字入力エリアAについて属性を調べると、その属性は「発行日」であることがわかる。

【0042】次に、変換テーブル22を参照し、抽出した文字列の前後に、調べた結果得られた属性に対応するタグを付加する処理を行なう（ステップS34）。即ち、タグ付加部26が、ステップS32で調べた結果として得られた属性に対応するタグを変換テーブル22より読み出して、ステップ30で抽出した文字列の前後に、そのタグを付加する。こうして得られたタグ付きの文字列はタグ付き文書30の一部として出力される。

【0043】従って、例えば、変換テーブルの内容が図10に示した如くであり、調べた結果得られた属性が「発行日」である場合、タグとしては「<date></date>」が読み出される。そして、読み出されたタグは、ステップ30で抽出された文字列「〇年〇月〇日」の前後に付加され、その結果、タグ付きの文字列「<date>〇年〇月〇日</date>」が得られ、これがタグ付き文書の一部となる。

【0044】次に、作成された定型文書28中に、文字列を抽出すべき文字入力エリアが残っているか否かを判定する処理を行なう（ステップS36）。即ち、文字列抽出部24が、定型文書28内の文字入力エリアのうち、未だ文字列の抽出されていない文字入力エリアが存在するかどうかを判定する。そして、存在する場合にはステップ30に戻って次の座標を読み出す処理を行ない、以下、前述と同様の処理を繰り返す。存在しない場合には、図11に示す処理ルーチンを終了する。

【0045】このようにして、本実施例においても、第1の実施例と同様にタグ付き文書を作成することができる。即ち、図10に示す変換テーブル22の内容に従って、図6の定型文書内の所定の文字列にタグが付された場合は、前述の図8に示すようなタグ付き文書が作成される。

【0046】以上述べたように、本実施例によれば、第1の実施例と同様に、文書作成者がタグの意味や使い方或いは文書の論理構造などを理解していなくても、容易にタグ付き文書を作成することができる。また、前述したように、文字入力エリアが異なっている、同一の属性を設定することは可能であるので、異なる文字入力エ

リアの文字列に同一のタグを付加したいような場合でも、第1の実施例に比較して簡単に付加することができる。

【0047】なお、本実施例においては、図11のフローチャートに示すように、まず、文字入力エリアより文字列を抽出し、次に、その文字入力エリアの属性を調べるようにしているが、第1の実施例における図7のフローチャートに類似するように、まず、変換テーブルより属性を読み出し、その属性に該当する文字入力エリアより文字列を抽出するようにしても良い。一般に、市販されているDTPソフトなどでは、文書について属性を指定すると、その属性が設定されている文字入力エリアより文字列を読み出すことができるようになっているものが多いので、そのようなソフトを利用することにより、上記のような処理は容易に実現することができる。

【0048】なお、本発明は上記した実施例や実施形態に限られるものではなく、その要旨を逸脱しない範囲において種々の態様にて実施することが可能であることは言うまでもない。

【0049】即ち、上記した実施例では、変換テーブル22の内容として、文字入力エリアの座標や属性を用いていたが、本発明はこれらに限定されるものではなく、文字入力エリアに関連する情報であればどのようなものでも良い。例えば、各記入項目の先頭に付された番号や記号やタイトルなども関連情報として利用することができる。

【0050】また、上記した実施例では、タグを抽出した文字列の前後に付加していたが、作成するタグ付き文書の構成によっては、タグを文字列の先頭のみあるいは後尾のみに付加しても良く、また、文字列の途中の特定の位置に付加しても良い。

【図面の簡単な説明】

【図1】本発明の第1の実施例としてのタグ付き文書作成装置を示すブロック図である。

【図2】タグ付き文書を作成するまでの一連の処理手順を示すフローチャートである。

【図3】作成された定型フォーマットの一例を示す説明図である。

【図4】図3の定型フォーマットにおける各文字入力エリアA～Eの座標を示す説明図である。

【図5】作成された変換テーブルの内容の一例を示す説明図である。

【図6】図3の定型フォーマットを用いて文書を入力した状態を示す説明図である。

【図7】図2のタグ付き文書作成処理ステップS16の詳細処理ルーチンを示すフローチャートである。

【図8】図6の定型文書について作成されたタグ付き文書の一例を示す説明図である。

【図9】図3の定型フォーマットにおける各文字入力エリアA～Eに対して設定された属性の一例を示す説明図

である。

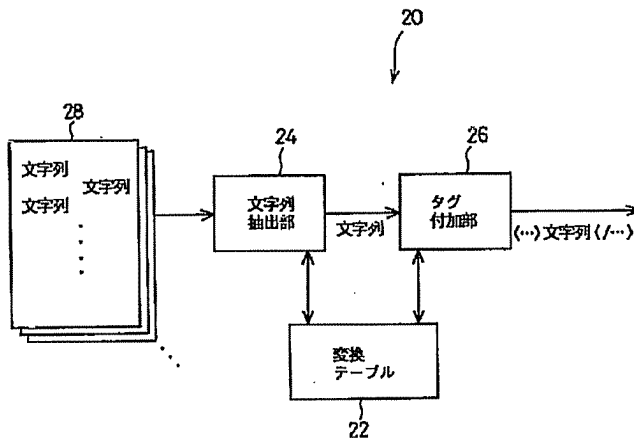
【図10】作成された変換テーブルの内容の一例を示す説明図である。

【図11】本発明の第2の実施例におけるタグ付き文書作成処理ステップS16の詳細処理ルーチンを示すフローチャートである。

【符号の説明】

20…文書作成装置
22…変換テーブル
24…文字列抽出部
26…タグ付加部
28…入力文書
30…タグ付き文書
A～E…文字入力エリア

【図1】



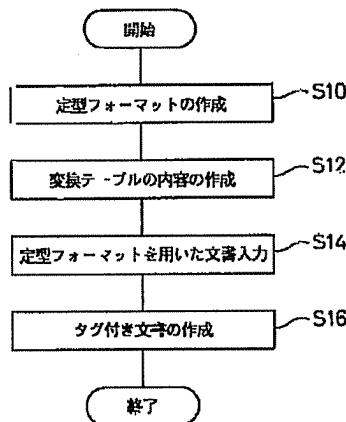
【図5】

座 標	付加すべきタグ
(146, 25)	→ <date> </date>
(16, 41)	→ <title> </title>
(16, 61)	→ <term> </term>
.	.
.	.
.	.
.	.
.	.
(16, 246)	→ <tech> </tech>
(16, 256)	→ <keyword> </keyword>

【図8】

<date> 〇年〇月〇日</date>
<title> ××××××××</title>
<term> *年*月*日～*月*日</term>
.
.
.
.
.
<tech> □□□□□□</tech>
<keyword> ◇◇◇◇◇◇</keyword>

【図2】



【図3】

技術報告書		A 発行日:
B [件名:]		
C [期間:]	報告者	所属:
メンバー:	氏名:	
目的:		
概要および結論:		
特 許 出 願: 出願済 (No.)・依 頼 中・予 定・未 定・し ない		
創 意 工 夫 案: 提案済・予 定・未 定・し ない		
[基 礎 分 野:] D		
[トピック:] E		
備考:		

【図4】

技術報告書

発行日: (146,25)

件名: (16,41)

期間: (16,61)

メンバー:

目的:

概要および結論:

特許出願: 出願済 (No.)・依頼中・予定・未定・しない
 創発工実案: 提案済・予定・未定・しない
 技術分類: D
 キーワード: E
 備考:

【図6】

技術報告書

発行日: ○年○月○日

件名: ××××××

期間: ×年×月×日～×月×日

メンバー:

目的:

概要および結論:

特許出願: 出願済 (No.)・依頼中・予定・未定・しない
 創発工実案: 提案済・予定・未定・しない
 技術分類: □□□□□～I
 キーワード: ◇◇◇◇◇～E
 備考:

【図9】

技術報告書

発行日:

属性: 件名

属性: 期間

メンバー:

目的:

概要および結論:

特許出願: 出願済 (No.)・依頼中・予定・未定・しない
 創発工実案: 提案済・予定・未定・しない
 技術分類: D
 キーワード: E
 備考: 属性: キーワード

【図10】

属性 付加すべきタグ

発行日 → <date> </date>

件名 → <title> </title>

期間 → <term> </term>

・

・

・

・

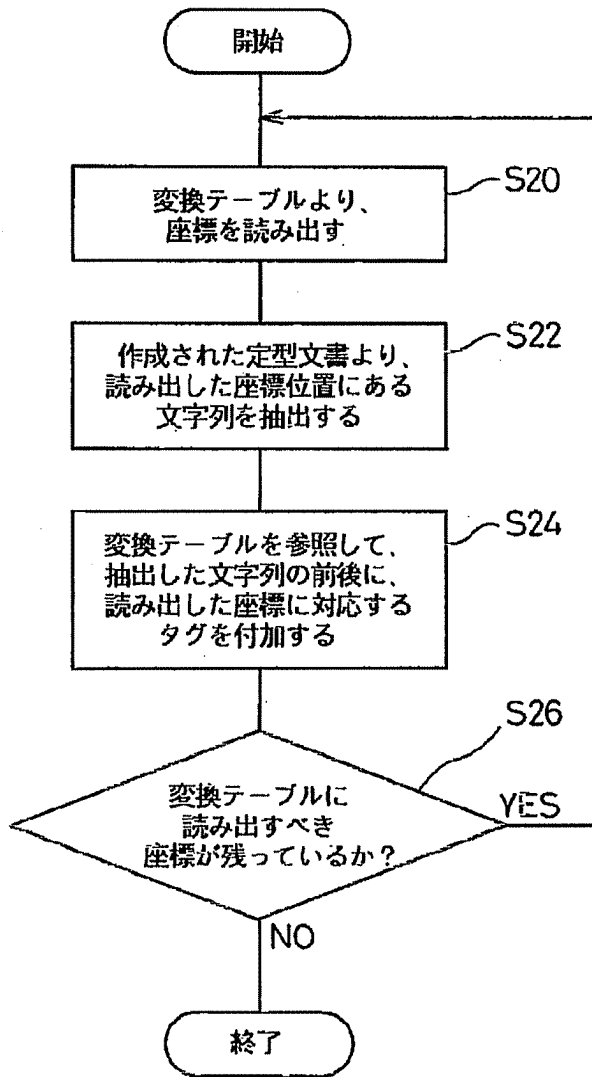
・

・

技術分類 ▶ <tech> </tech>

キーワード ▶ <keyword> </keyword>

【図7】



【図11】

